(51) International Patent Classification: C07H 21/04, C12Q 1/68	**A1**	(11) International Publication Number: **WO 00/69882**
		(43) International Publication Date: 23 November 2000 (23.11.2000)

(21) International Application Number: PCT/US00/13168	**Published**
(22) International Filing Date: 15 May 2000 (15.05.2000)	
(30) Priority Data: 60/134,180 13 May 1999 (13.05.1999) US	
(60) Parent Application or Grant THE PENN STATE RESEARCH FOUNDATION [/]; (). GREGER, Douglas, L. [/]; (). GREGER, Douglas, L. [/]; (). RIGAUT, Kathleen, D. ; ().	

(54) Title: GENETIC MARKER FOR MEAT QUALITY, GROWTH, CARCASS AND REPRODUCTIVE TRAITS IN LIVESTOCK

(54) Titre: MARQUEUR GENETIQUE POUR LA QUALITE DE VIANDE, LA CROISSANCE, LA CARCASSE ET LES CARACTERISTIQUES REPRODUCTRICES DU BETAIL

(57) Abstract

Compositions and methods are provided for identifying polymorphisms associated with growth and reproductive traits in livestock.

(57) Abrégé

L'invention concerne des compositions et des procédés permettant l'identification des polymorphismes associés à la croissance et aux caractéristiques de reproduction du bétail.

PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [7] : C07H 21/04, C12Q 1/68	**A1**	(11) International Publication Number: **WO 00/69882** (43) International Publication Date: 23 November 2000 (23.11.00)

(21) International Application Number: PCT/US00/13168

(22) International Filing Date: 15 May 2000 (15.05.00)

(30) Priority Data:
60/134,180   13 May 1999 (13.05.99)   US

(71) Applicant (for all designated States except US): THE PENN STATE RESEARCH FOUNDATION [US/US]; 304 Old Main, University Park, PA 16802-7000 (US).

(72) Inventor; and
(75) Inventor/Applicant (for US only): GREGER, Douglas, L. [US/US]; 23 Coventry Lane, State College, PA 16803 (US).

(74) Agent: RIGAUT, Kathleen, D.; Dann, Dorfman, Herrell and Skillman, Suite 720, 1601 Market Street, Philadelphia, PA 19103 (US).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
 With international search report.

(54) Title: GENETIC MARKER FOR MEAT QUALITY, GROWTH, CARCASS AND REPRODUCTIVE TRAITS IN LIVESTOCK

(57) Abstract

Compositions and methods are provided for identifying polymorphisms associated with growth and reproductive traits in livestock.

Description

5

10

15

20

25

30

35

40

45

50

55

GENETIC MARKER FOR MEAT QUALITY, GROWTH, CARCASS AND REPRODUCTIVE TRAITS IN LIVESTOCK

FIELD OF THE INVENTION

This invention relates generally to the detection of genetic differences associated with growth, body composition and reproductive traits among livestock. More specifically, the invention provides compositions and methods for predicting heritability of certain traits related to steroid biosynthesis and metabolism.

BACKGROUND OF THE INVENTION

Several publications are referenced in this application by author name, year and journal of publication in parentheses in order to more fully describe the state of the art to which this invention pertains. The disclosure of each of these publications is incorporated by reference herein.

Steroid hormones play a crucial role in the differentiation, development, growth and physiological function of most animal tissues. The first and rate-limiting step in the biosynthesis of all steroid hormones is the conversion of cholesterol into pregnenolone by the cholesterol side chain cleavage enzyme p450scc. The gene which encodes P450scc is termed CYP11a1. Cytochromes P450 are a diverse group of heme-containing mono-oxygenases (termed CYP's; see Nelson et al., DNA Cell Biol. (1993) 12: 1-51) that catalyze a variety of oxidative conversions, notably of steroids but also of fatty acids and xenobiotics. CYP's are most abundantly expressed in the testis, ovary, placenta, adrenal glands and liver. In the reproductive organs, such as testis, ovary and placenta, the most

important steroid hormones produced are the androgens (e.g., testosterone), the estrogens (e.g., estradiol) and progestins (e.g., progesterone). In the adrenal glands, the most important steroids are the
5   mineralcorticoids (e.g., aldosterone) and the glucocorticoids (e.g., cortisol).

The frequent occurrence of off-odors or off-tastes in cooked pork from boars, commonly known as "boar odor" or "boar taint", is the primary reason for the common
10   practice of castration in swine production.   5α-androstenone (5α-androst-16-en-3-one), an important compound responsible for boar taint, is synthesized in the boar testis along with other 16-androstene steroids, androgens, and estrogens.   At puberty, testicular
15   production of Δ16-androstenes, in particular 5α-androstenone (androstenone), increases sharply. This results in the accumulation of androstenone in various body compartments, notably in fat deposits throughout the body and in the submaxillary salivary gland (SMG),
20   where there is a specific binding protein for Δ16-androstenes. Concentration of androstenone and other Δ16-androstenes in the SMG are highly correlated with concentrations of Δ16- androstenes in the fat. Measurement of Δ16-androstenes in the SMG is used, in
25   fact, as a test method to determine the presence or absence of boar taint. Thus, due to this increase in Δ16-androstenes, it is common in the industry to castrate the young male boars to minimize this taint in the meat.   However, if the problem of boar taint were
30   overcome, raising boars rather than raising castrates (barrows) for pork would have considerable economic advantages. Although boars and barrows gain weight at equivalent rates, boars produce carcasses containing 20-30% less fat. Thus, boars are much more efficient at
35   producing lean muscle. In addition, boars utilize feed

2

more efficiently than barrows (10% less feed consumed per unit of body weight). Since feed represents the major cost in swine production, raising boars for pork would have significant economic advantages.

In the United States, approximately 90 million hogs are slaughtered annually with an approximate value of $11 billion. Feed accounts for the major portion of the costs of swine production, accounting for roughly 70% of production costs. Thus, a 10% improvement in feed efficiency would produce savings of 7% of the total cost of production. On a nation-wide basis, considering male swine only, this translates to total market savings of $335 million. The loss of production efficiency caused by the practice of castration represents a very large economic loss to the swine industry throughout the world.

Identification of the inheritance pattern(s) and genetic bases for alterations in steroid biosynthesis in livestock has utility in the production of meat, dairy and egg products of higher quality. It is an object of the present invention to provide compositions and methods for identifying such genetic alterations.

SUMMARY OF THE INVENTION

In accordance with the present invention, methods for identifying genetic alterations associated with steroid biosynthesis are provided. In one embodiment of the invention, the presence or absence of a polymorphic marker in the CYP11a1 DNA of a test subject is determined. Such test subjects are selected from important livestock species, including without limitation, pigs, cows, chickens and sheep. In accordance with the present invention, it has been determined that certain polymorphisms in the CYP11a1 gene are associated with increased growth, reproductive

3

and carcass traits. Thus, screening methods are provided for identifying those test subjects which possess these beneficial CYP11al alleles. Identification of such livestock facilitates the

5   implementation of breeding programs for developing stock having these improved genetic traits.

As is well known to those of skill in the art, a variety of techniques may be utilized when comparing nucleic acid molecules for sequence differences. These

10   include by way of example, restriction fragment length polymorphism analysis, heteroduplex analysis, single strand conformation polymorphism analysis, denaturing gradient electrophoresis and temperature gradient electrophoresis.

15   In a preferred embodiment of the invention, the CYP11al polymorphism is a restriction fragment polymorphism and the assay comprises identifying the CYP11al gene from genetic material isolated from the test subject; exposing the gene to a restriction enzyme

20   that yields restriction fragments of the gene of varying length; separating the restriction fragments to form a restriction pattern, such as by electrophoresis or HPLC separation; and comparing the resulting restriction fragment pattern from a test subject CYP11al gene that

25   is either known to have or not to have the desired marker. If a test subject tests positive for the marker, such a subject can be considered for inclusion in the breeding program. If the test subject does not test positive for the marker genotype, the test subject can

30   be culled from the group and otherwise used.

In a particularly preferred embodiment, the test subject is a pig, the polymorphism is in the 5'UTR of the CYP11al gene and the restriction enzyme is SphI. Thus, in this aspect, it is an object of the invention

35   to provide a method of screening pigs to determine those

4

more likely to have decreased testis weight and reduced
boar taint, longer carcasses, improved rate of gain, or
heavier weaning weights when bred to or to select
against pigs which have alleles indicating larger testis
size, increased boar taint, reduced carcass length,
lower rate of gain, or lighter weaning weights. As used
herein "smaller testis size" means a significant
decrease in testis size below the mean for a given
population. As used herein "reduced boar taint" means a
significant decrease in boar taint below the mean for a
given population. As used herein "increased carcass
length" means a significant increase in carcass length
above the mean for a given population. As used herein
"higher rate of gain" means a significant increase in
rate of gain above the mean for a given population. As
used herein "heavier weaning weights" mean an increase
in weaning weight above the mean for a given population.
The method of the invention comprises the steps: 1)
obtaining a sample of genomic DNA from a pig; and
2) analyzing the genomic DNA obtained in 1) to determine
which CYP11a1 allele(s) is/are present. Briefly, a
sample of genetic material is obtained from a pig, and
the sample is analyzed to determine the presence or
absence of a polymorphism in the CYP11a1 gene that is
correlated with reduced boar taint, smaller testis size,
increased carcass length, higher rate of gain, and/or
increased weaning weight.

In a most preferred embodiment the gene is isolated
by the used of primers and DNA polymerase to amplify a
specific region of the gene which contains the
polymorphism. Next the amplified region is digested with
a restriction enzyme and fragments are separated.
Visualization of the RFLP pattern is by simple staining
of the fragments, or by labeling the primers or the
nucleoside triphosphates used in amplification.

5

In another embodiment, the invention comprises a
method for identifying a genetic marker for boar taint,
testis size, carcass length, rate of gain, and/or
weaning weight in a particular population. Male and
5   female pigs of the same breed or breed cross or similar
genetic lineage are bred, and traits such as boar taint,
testis size, carcass length, rate of gain, and/or
weaning weight are determined. A polymorphism in the
CYP11a1 gene of each pig is identified and associated
10   with the traits of boar taint, testis size, carcass
length, rate of gain, and/or weaning weight. Preferably,
RFLP analysis is used to determine the polymorphism, and
most preferably, the DNA is digested with the
restriction endonuclease SphI, or other restriction
15   endonuclease that differentially cleaves the restriction
site based on the presence or absence of the
polymorphism.

Methods are also provided to establish linkage
between specific alleles of alternative DNA markers and
20   alleles of DNA markers known to be associated with a
particular gene (e.g. the CYP11a1 gene discussed
herein), which have been previously shown to be
associated with a particular trait. Thus, selection for
pigs likely to have reduced boar taint, smaller testes,
25   increased carcass length, higher rate of gain, and/or
heavier weaning weights, or alternatively to select
against pigs likely to have increased boar taint, larger
testes, reduced carcass length, lower rate of gain,
and/or lighter weaning weights, may be done indirectly,
30   by selecting for certain alleles of a CYP11a1 associated
marker through the selection of specific alleles of
alternative markers located on the same chromosome as
CYP11a1.

The invention further comprises kits for evaluating
35   a sample of test subject DNA for the presence in test

6

subject genetic material of a desired marker located in
the test subject CYP11a1 gene indicative of the
inheritable traits of boar taint (in the pig), testis
size, carcass length, rate of gain, and/or weaning
weight. At a minimum, using the pig as the test subject,
the kit is a container with one or more reagents that
identify a polymorphism in the pig CYP11a1 gene.
Preferably, the reagent is a set of oligonucleotide
primers capable of amplifying a fragment of the pig
CYP11a1 gene that contains the polymorphism. More
preferably, the kit further contains a restriction
enzyme that cleaves the pig CYP11a1 gene in at least one
place. In a most preferred embodiment the restriction
enzyme is SphI or one which cuts at the same recognition
site.

The following definitions are provided to
facilitate an understanding of the present invention:

The term "corresponds to" is used herein to mean
that a polynucleotide sequence is homologous to all or a
portion of a reference polynucleotide sequence, or that
a polypeptide sequence is identical to a reference
polypeptide sequence. In contradistinction, the term
"complementary to" is used herein to mean that the
complementary sequence is homologous to all or a portion
of a reference polynucleotide sequence. For
illustration, the nucleotide sequence "TATAC"
corresponds to a reference sequence "TATAC" and is
complementary to a reference sequence "GTATA".
Hybridization probes may be DNA or RNA, or any synthetic
nucleotide structure capable of binding in a base-
specific manner to a complementary strand of nucleic
acid. For example, probes include peptide nucleic acids,
as described in Nielsen et al., Science 254:1497-1500
(1991).

7

"Linkage" describes the tendency of genes, alleles, loci or genetic markers to be inherited together as a result of their location on the same chromosome, and is measured by percent recombination (also called recombination fraction, or θ) between the two genes, alleles, loci or genetic markers. The closer two loci physically are on the chromosome, the lower the recombination fraction will be. Normally, when a polymorphic site from within a disease-causing gene is tested for linkage with the disease, the recombination fraction will be zero, indicating that the disease and the disease-causing gene are always co-inherited. In rare cases, when a gene spans a very large segment of the genome, it may be possible to observe recombination between polymorphic sites on one end of the gene and causitive mutations on the other. However, if the causative mutation is the polymorphism being tested for linkage with the disease, no recombination will be observed.

"Centimorgan" is a unit of genetic distance signifying linkage between two genetic markers, alleles, genes or loci, corresponding to a probability of recombination between the two markers or loci of 1% for any meiotic event.

"Linkage disequilibrium" or "allelic association" means the preferential association of a particular allele, locus, gene or genetic marker with a specific allele, locus, gene or genetic marker at a nearby chromosomal location more frequently than expected by chance for any particular allele frequency in the population.

An "oligonucleotide" can be DNA or RNA, and single-

8

or double-stranded. Oligonucleotides can be naturally
occurring or synthetic, but are typically prepared by
synthetic means.

5   The term "primer" refers to an oligonucleotide,
capable of acting as a point of initiation of DNA
synthesis under conditions in which synthesis of a
primer extension product complementary to a nucleic acid
strand is induced, i.e., in the presence of four
10   different nucleoside triphosphates and an agent for
polymerization (i.e., DNA polymerase or reverse
transcriptase) in an appropriate buffer and at a
suitable temperature. A primer is preferably a single-
stranded oligonucleotide. The appropriate length of a
15   primer depends on the intended use of the primer but
typically ranges from 15 to 30 nucleotides. Short primer
molecules generally require cooler temperatures to form
sufficiently stable hybrid complexes with the template.
A primer need not reflect the exact sequence of the
20   template but must be sufficiently complementary to
hybridize with a template. The term "primer" may refer
to more than one primer, particularly in the case where
there is some ambiguity in the information regarding one
or both ends of the target region to be amplified. For
25   instance, if a region shows significant levels of
polymorphism or mutation in a population, mixtures of
primers can be prepared that will amplify alternate
sequences. A primer can be labeled, if desired, by
incorporating a label detectable by spectroscopic,
30   photochemical, biochemical, immunochemical, or chemical
means. For example, useful labels include ^{32}P,
fluorescent dyes, electron-dense reagents, enzymes (as
commonly used in an ELISA), biotin, or haptens and
proteins for which antisera or monoclonal antibodies are
35   available. A label can also be used to "capture" the

9

primer, so as to facilitate the immobilization of either the primer or a primer extension product, such as amplified DNA, on a solid support.

"Chromosome 7 set" in boars for example, means the two copies of chromosome 7 found in somatic cells or the one copy in germ line cells of a test subject or family member. The two copies of chromosome 7 may be the same or different at any particular allele, including alleles at or near the locus of interest. The chromosome 7 set may include portions of chromosome 7 collected in chromosome 7 libraries, such as plasmid, yeast, or phage libraries, as described in Sambrook et al., Molecular Cloning, 2nd Edition, and in Mandel et al., Science 258:103-108 (1992).

"Penetrance" is the percentage of individuals with a defective gene or polymorphism who show some symptoms of a trait resulting from that genetic alteration. Expressivity refers to the degree of expression of the trait (e.g., mild, moderate or severe).

"Polymorphism" refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. A polymorphic marker is the locus at which divergence occurs. Preferred markers have at least two alleles, each occurring at frequency of greater than 1%. A polymorphic locus may be as small as one base pair difference. Polymorphic markers suitable for use in the invention include restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, and other microsatellite sequences.

"Restriction fragment length polymorphism" (RFLP) means a variation in DNA sequence that alters the length of a restriction fragment as described in Botstein et al., Am. J. Hum. Genet. 32:314-331 (1980). The restriction fragment length polymorphism may create or delete a restriction site, thus changing the length of the restriction fragment. For example, the DNA sequence GAATTC are the six bases, together with its complementary strand CTTAAG which comprises the recognition and cleavage site of the restriction enzyme EcoRI. Replacement of any of the six nucleotides on either strand of DNA to a different nucleotide destroys the EcoRI site. This RFLP can be detected by, for example, amplification of a target sequence including the polymorphism, digestion of the amplified sequence with EcoRI, and size fractionation of the reaction products on an agarose or acrylamide gel. If the only EcoRI restriction enzyme site within the amplified sequence is the polymorphic site, the target sequences comprising the restriction site will show two fragments of predetermined size, based on the length of the amplified sequence. Target sequences without the restriction enzyme site will only show one fragment, of the length of the amplified sequence. Similarly, the RFLP can be detected by probing an EcoRI digest of Southern blotted DNA with a probe from a nearby region such that the presence or absence of the appropriately sized EcoRI fragment may be observed. RFLP's may be caused by point mutations which create or destroy a restriction enzyme site, VNTR's, dinucleotide repeats, deletions, duplications, or any other sequence-based variation that creates or deletes a restriction enzyme site, or alters the size of a restriction fragment.

"Variable number of tandem repeats" (VNTR's) are

11

short sequences of nucleic acids arranged in a head to
tail fashion in a tandem array, and found in each
individual, as described in Wyman et al., Proc. Nat.
Acad. Sci. 77:6754-6758 (1980). Generally, the VNTR

5   sequences are comprised of a core sequence of at least
16 base pairs, with a variable number of repeats of that
sequence. Additionally, there may be variation within
the core sequence, Jefferys et al., Nature 314:67-72
(1985). These sequences are highly individual, and

10   perhaps unique to each individual. Thus, VNTR's may
generate restriction fragment length polymorphisms, and
may additionally serve as size-based amplification
product differentiation markers.

15   "Microsatellite sequences" comprise segments of at
least about 10 base pairs of DNA consisting of a
variable number of tandem repeats of short (1-6 base
pairs) sequences of DNA(Clemens et al., Am. J. Hum.
Genet. 49:951-960 1991). Microsatellite sequences are

20   generally spread throughout the chromosomal DNA of an
individual. The number of repeats in any particular
tandem array varies greatly from individual to
individual, and thus, microsatellite sequences may serve
to generate restriction fragment length polymorphisms,

25   and may additionally serve as size-based amplification
product differentiation markers.

BRIEF DESCRIPTION OF THE DRAWINGS

30   Figure 1 depicts the sequence of approximately 630
base pairs of the 5'untranslated region of the porcine
CYP11A1 gene (SEQ ID NO: 1). The PCR fragment was
produced using DNA extracted from porcine testis
samples. The primers used were forward primer (SEQ ID

35   NO:2) and reverse primer (SEQ ID NO:3).

12

Figure 2 depicts the polymorphic pattern of Sphl
digested PCR product. The forward and reverse primers
were used in the following PCR conditions: Two minutes @
94°C, 35 cycles of one minute @ 94°C, one minute @ 55°C,
5   one minute @ 72°C and a final two minutes @ 72°C.
Samples were digested with SphI (New England Biolabs)
and separated on 1.5% agarose gel at 50 volts for 45
minutes at room temperature. Gels were stained with
ethidium bromide. Lane 1: low molecular weight
10   markers; Lane 2: undigested PCR fragment; Lanes 3 and 7:
genotype CT; and Lanes 4-6: genotype CC. A Restriction
Fragment Length Polymorphism (RFLP) was discovered
whereby the 630 bp PCR fragment from CC pigs was
digested into a 450 bp product while the PCR fragment
15   from the CT pigs was only partially digested, which
indicates the presence of the T allele.

Figure 3 depicts the concentrations of submaxillary
salivary gland (SMG) Δ-16 androstenes in boars of the CC
20   versus the CT genotype. Five out of thirty of the CC
boars exhibited SMG Δ-16 androstene concentrations
greater than the recommended threshold level for
identifying tainted carcasses (55 µg/g SMG). All of
boars carrying the T allele (n=20) were below the
25   recommended threshold level for boar taint.

Figure 4 is a table that shows the observed
differences in various growth, carcass, and reproductive
traits of CC versus CT boars. The greater weights of
30   testes, submaxillary glands and bulbourethral glands, as
well as higher concentrations of SMG Δ-16-androstenes,
are all indications of higher boar taint in the CC
boars. Surprisingly the CC boars also had 5.9% increase
in rate of gain and longer carcasses as well.

35

13

Figure 5 shows the sequence of the bovine CYP11a1 gene, including 948 nucleotide of the 5' UTR.

Figure 6 shows the sequence of the chicken CYP11a1 gene, including 137 nucleotide of the 5'UTR.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, materials and methods are provided for diagnosing genetic alterations in the CYP11a1 gene associated with aberrant or increased sterioid biosynthesis in livestock. In the mouse, polymorphic variation in CYP11a1 is responsible for genetic differences in testosterone production. In mouse, CYP11a1 maps to chromosome 9. This region is syntenic with porcine chromosome 7.

A principle cause of taint in the boar is the presence of the Δ-16 steroid, androstenone, which is one of many steroids produced in the boar testis. Androstenone and androstenone metabolites such as androstenol are secreted by the testis and sequestered in the submaxillary salivary glands (SMG). During mating behavior these steroids are released into the air through the saliva and function as sexual pheromones whereby they induce estrous behavior in female pigs (sows). Since Δ-16 steroids are highly lipophilic, androstenone is also stored in body fat, where its presence in high concentrations contributes to the off-flavors in pork known as boar taint.

Concentrations of androstenone in the fat are highly heritable. A quantitative trait locus (QTL) has been identified for fat androstenone (microsatellite marker SO102), which is located on porcine chromosome 7 in the region of the swine leukocyte antigen complex (SLA). In accordance with the present invention, a particular genetic polymorphic sequence has been

14

identified which is associated with androstenone
production and boar taint.

5   The presence of a quantitative trait locus (QTL)
for fat androstenone on chromosome 7 in the pig suggests
that porcine CYP11a1 may be located on chromosome 7 and,
as the rate limiting enzyme in steroid synthesis may be
an important control point for androsterone synthesis
and the occurrence of boar taint.

10   A genomic search was conducted to compare 2.4 kb of
the untranslated region (5`UTR) of the porcine CYP11a1
·gene from a preselected group of boars in order to
determine if polymorphisms exist which are associated
with compounds which cause boar taint. First,
comparisons of the genotypes of five "high taint" and
15   five "low taint" boars by direct sequencing of PCR
products (using the ABI Prism 377 at the Nucleic Acid
Facility, Penn State University Biotechnology Institute)
revealed the presence of one single nucleotide
polymorphism (SNP) in the entire 2.4 kb 5' UTR. This
20   SNP (CT allele) was discovered only in boars that
exhibited low concentrations of delta-16 steroids in the
salivary gland, a measurement that is highly correlated
with androstenone concentrations in the fat. This
polymorphism consists of either a thymidine (T) or a
25   cytosine (C) at position - 155 from the start site of
translation. The polymorphism was located in a
restriction enzyme recognition site such that the
presence of the T allele would change the restriction
fragment length pattern observed after digestion with
30   specific restriction enzymes. In this particular case,
the restriction enzyme used was Sph1 (New England
Biolabs). Additional restriction enzymes are available
which are able to cut the same DNA sequence. Presence or
absence of the T allele was determined by examination of
35   restriction digests of CYP11a1 5'UTR using Sph1.

15

Presence of the T allele, either homozygous (TT) or
heterozygous (CT), was associated with low boar taint.
Presence of the CC allele was associated with high boar
taint, as well as with increased testis weight,
bulbourethral gland length and weight and submaxillary
salivary gland weights. In addition, boars that
possessed the CC allele exhibited a 5.9% improvement in
rate of gain as well as longer carcasses.

The discovery that this polymorphism is associated
with increased rate of gain and carcass length in
addition to its effects on reproductive traits indicates
that this polymorphism affects many other growth and
developmental traits. Thus, presence or absence of this
polymorphism may also be associated with feed efficiency
and with birth weight. The association of this
polymorphism with reproductive traits such as testis
weight, bulbourethral gland length and weight,
submaxillary gland weight, and Δ-16 steroid
concentrations, are all indications of a general effect
on gonadal steroid production.

The data presented herein indicate that the
presence or absence of the CYP11a1 polymorphism may have
effects on other reproductive traits such as ovulation
rate, litter size, milk production, and fertility (both
male and female). Additionally, since the adrenal gland
is another site where CYP11a1 is expressed to produce
glucocorticoid steroids such cortisol, this polymorphism
may be associated with disease response traits since
these traits are known to be modulated by adrenal
steroids.

In a further aspect of the invention, this genetic
marker may also be used in combination with other
genetic markers to produce favorable combinations of
alleles or to select against those test subjects
carrying unfavorable combinations. Examples of some of

16

these previously identified genes are: tumor necrosis
factor alpha (TNFa), CYP11a1, prolactin (PRL), estrogen
receptor (ER) and prolactin receptor (PRIR). Examples
of some of these previously identified microsatellite
5   markers are: S0064, S0102, S0078, S0158, S0066, SW304,
SW1083, S0101, and S0212.

       Additional polymorphisms in the porcine CYP11a1
gene may be identified using the methods of the present
invention. Such alterations may occur in the
10   untranslated region of the gene but may also be
identified in the translated region, as well as in the
intronic and exonic sequences. It is likely that a
subset of these changes will cause or be associated with
changes in androgen function and phenotypic traits.
15   Once such genetic alterations are identified, it is
possible to introduce these or similar changes into the
genome by known techniques in order to produce
transgenic animals that possess a desired CYP11a1
genotype. The data further suggest that polymorphisms
20   in homologous areas of CYP11a1 of other agriculturally
important species are likely to cause or be associated
with similar changes in function and phenotype.

       In a further aspect of the invention, the
corresponding CYP11a1 sequences from the cow and the
25   chicken are provided. This information facilitates
genomic scanning of the 5'UTR of the bovine or chicken
CYP11a1 to reveal polymorphisms that are associated with
growth, carcass traits, and reproduction (including milk
production and egg production).

30
       DIAGNOSTIC KITS FOR PRACTICING
       THE METHODS OF THE INVENTION
       The present invention also includes kits for the
practice of the methods of the invention. The kits
35   comprise a vial, tube, or any other container which

17

contains one or more oligonucleotides, which hybridizes
to a DNA segment which DNA segment which is or is linked
to the CYP11a1 gene. Some kits contain two such
oligonucleotides, which serve as primers to amplify a
5   segment of chromosome DNA. The segment selected for
amplification can be a CYP11a1 gene that includes a site
at which a variation is known to occur. Some kits
contain a pair of oligonucleotides for detecting
precharacterized variations. For example, some kits
10   contain oligonucleotides suitable for allele-specific
oligonucleotide hybridization, or allele-specific
amplification hybridization. The kits of the invention
may also contain components of the amplification system,
including PCR reaction materials such as buffers and a
15   thermostable polymerase. In other embodiments, the kit
of the present invention can be used in conjunction with
commercially available amplification kits, such as may
be obtained from GIBCO BRL (Gaithersburg, Md.)
Stratagene (La Jolla, Calif.), Invitrogen (San Diego,
20   Calif.), Schleicher & Schuell (Keene, N.H.), Boehringer
Mannheim (Indianapolis, Ind.). The kits may optionally
include positive or negative control reactions or
markers, molecular weight size markers for gel
electrophoresis, and the like. The kits usually include
25   labeling or instructions indicating the suitability of
the kits for diagnosing steroid biosynthesis alterations
and indicating how the oligonucleotides are to be used
for that purpose. The term "label" is used generically
to encompass any written or recorded material that is
30   attached to, or otherwise accompanies the diagnostic at
any time during its manufacture, transport, sale or use.

MODES OF PRACTICING THE INVENTION

1. Linkage Analysis
35   Determining linkage between a polymorphic marker

18

and a locus associated with a particular phenotype is performed by mapping polymorphic markers and observing whether they co-segregate with the high taint phenotype (for example) on a chromosome in an informative meiosis.

5   See, e.g., Kerem et al., Science 245:1073-1080 (1989); Monaco et al., Nature 316:842 (1985); Yamoka et al., Neurology 40:222-226 (1990), and as reviewed in Rossiter et al., FASEB Journal 5:21-27 (1991). A single pedigree rarely contains enough informative meioses to provide

10   definitive linkage, because families are often small and markers may be not sufficiently informative. For example, a marker may not be polymorphic in a particular family.

Linkage may be established by an affected sib-pairs

15   analysis as described in Terwilliger & Ott, Handbook of Human Genetic Linkage (Johns Hopkins, Md., 1994), Ch. 26. This approach requires no assumptions to be made concerning penetrance or variant frequency, but only takes into account the data of a relatively small

20   proportion (i.e., the SIB pairs) of all the family members whose phenotype and polymorphic markers have been determined. Specifically, the affected SIB pairs analysis scores each pair of affected SIBS as sharing (concordant) or not sharing (discordant) the same

25   allelic variant of each polymorphic marker. For each marker, a probability is then calculated that the observed ratio of concordant to discordant SIB pairs would arise without linkage of the marker.

As described in Thompson & Thompson, Genetics in

30   Medicine, 5th ed, 1991, W.B. Saunders Company, Philadelphia, in linkage analysis, one calculates a series of likelihood ratios (relative odds) at various possible values of θ, ranging from $\theta = 0.0$ (no recombination) to $\theta = 0.50$ (random assortment). Thus, the

35   likelihood ratio at a given value of θ is (likelihood of

19

data if αloci are linked at θ)/(likelihood of data if
loci are unlinked). Evidence in support of linkage is
usually expressed as the $\log_{10}$ of this ratio and called a
"lod score" for "logarithm of the odds." For example, a
lod score of 5 indicates 100,000:1 odds that the linkage
being observed did not occur by chance.

The use of logarithms allows data collected from
different families to be combined by simple addition.
Computer programs are available for the calculation of
lod scores for differing values of θ. Available programs
include LIPED, and MLINK (Lathrop, Proc. Nat. Acad. Sci.
81:3443-3446 (1984).

For any particular lod score, a recombination
fraction may be determined from mathematical tables. See
Smith et al., Mathematical tables for research workers
in human.genetics (Churchill, London, 1961) and Smith,
Ann. Hum. Genet. 32:127-150 (1968). The value of θ at
which the lod score is the highest is considered to be
the best estimate of the recombination fraction, the
"maximum likelihood estimate".

Positive lod score values suggest that the two loci
are linked, whereas negative values suggest that linkage
is less likely (at that value of θ) than the possibility
that the two loci are unlinked. By convention, a
combined lod score of +3 or greater (equivalent to
greater than 1000:1 odds in favor of linkage) is
considered definitive evidence that two loci are linked.
Similarly, by convention, a negative lod score of -2 or
less is taken as definitive evidence against linkage of
the two loci being compared. If there are sufficient
negative linkage data, a locus can be excluded from an
entire chromosome, or a portion.thereof, a process
referred to as exclusion mapping. The search is then
focused on the remaining non-excluded chromosomal
locations. For a general discussion of lod scores and

20

linkage analysis, see, e.g., T. Strachan, Chapter 4,
"Mapping the human genome" in The Human Genome, 1992
BIOS Scientific Publishers Ltd. Oxford.

5   The data can also be subjected to haplotype
analysis. This analysis assigns allelic markers between
the chromosomes of an individual such that the number of
recombinational events needed to account for segregation
between generations is minimized. Linkage may also be
established by determining the relative likelihood of
10   obtaining observed segregation data for any two markers
when the two markers are located at a recombination
fraction θ, versus the situation in which the two
markers are not linked, and thus segregating
independently.

15

2. Isolation and Amplification of DNA

       Samples of patient, proband, test subject, or
family member genomic DNA are isolated from any
20   convenient source including saliva, buccal cells, hair
roots, blood, cord blood, amniotic fluid, interstitial
fluid, peritoneal fluid, chorionic villus, and any other
suitable cell or tissue sample with intact interphase
nuclei or metaphase cells. The cells can be obtained
25   from solid tissue as from a fresh or preserved organ or
from a tissue sample or biopsy. The sample can contain
compounds which are not naturally intermixed with the
biological material such as preservatives,
anticoagulants, buffers, fixatives, nutrients,
30   antibiotics, or the like.
       Methods for isolation of genomic DNA from these
various sources are described in, for example, Kirby,
DNA Fingerprinting, An Introduction, W.H. Freeman & Co.
New York (1992). Genomic DNA can also be isolated from
35   cultured primary or secondary cell cultures or from

21

transformed cell lines derived from any of the
aforementioned tissue samples. .

Samples of patient, proband, test subject or family
member RNA can also be used. RNA can be isolated from
5   tissues expressing the CYP11a1 gene as described in
Sambrook et al., supra. RNA can be total cellular RNA,
mRNA, poly A+ RNA, or any combination thereof. For best
results, the RNA is purified, but can also be unpurified
cytoplasmic RNA. RNA can be reverse transcribed to form
10   DNA which is then used as the amplification template,
such that the PCR indirectly amplifies a specific
population of RNA transcripts. See, e.g., Sambrook,
supra, Kawasaki et al., Chapter 8 in PCR Technology,
(1992) supra, and Berg et al., Hum. Genet. 85:655-658
15   (1990).

3. PCR Amplification

The most common means for amplification is
20   polymerase chain reaction (PCR), as described in U.S.
Pat. Nos. 4,683,195, 4,683,202, 4,965,188 each of which
is hereby incorporated by reference. If PCR is used to
amplify the target regions in blood cells, heparinized
whole blood should be drawn in a sealed vacuum tube kept
25   separated from other samples and handled with clean
gloves. For best results, blood should be processed
immediately after collection; if this is impossible, it
should be kept in a sealed container at 4° C until use.
Cells in other physiological fluids may also be assayed.
30   When using any of these fluids, the cells in the fluid
should be separated from the fluid component by
centrifugation.

Tissues should be roughly minced using a sterile,
disposable scalpel and a sterile needle (or two
35   scalpels) in a 5 mm Petri dish. Procedures for removing

22

paraffin from tissue sections are described in a variety
of specialized handbooks well known to those skilled in
the art.

5   To amplify a target nucleic acid sequence in a
sample by PCR, the sequence must be accessible to the
components of the amplification system. One method of
isolating target DNA is crude extraction which is useful
for relatively large samples. Briefly, mononuclear cells
from samples of blood, amniocytes from amniotic fluid,
10   cultured chorionic villus cells, or the like are
isolated by layering on sterile Ficoll-Hypaque gradient
by standard procedures. Interphase cells are collected
and washed three times in sterile phosphate buffered
saline before DNA extraction. If testing DNA from
15   peripheral blood lymphocytes, an osmotic shock
(treatment of the pellet for 10 sec with distilled
water) is suggested, followed by two additional washings
if residual red blood cells are visible following the
initial washes. This will prevent the inhibitory effect
20   of the heme group carried by hemoglobin on the PCR
reaction. If PCR testing is not performed immediately
after sample collection, aliquots of 10^6 cells can be
pelleted in sterile Eppendorf tubes and the dry pellet
frozen at -20° C until use.

25   The cells are resuspended (10^6 nucleated cells per
100 µl) in a buffer of 50 mM Tris-HCl (pH 8.3), 50 mM
KCl 1.5 mM MgCl$_2$, 0.5% Tween 20, 0.5% NP40 supplemented
with 100 µg/ml of proteinase K. After incubating at 56°
C for 2 hr, the cells are heated to 95° C for 10 min to
30   inactivate the proteinase K and immediately moved to wet
ice (snap-cool). If gross aggregates are present,
another cycle of digestion in the same buffer should be
undertaken. Ten µl of this extract is used for
amplification.

35   When extracting DNA from tissues, e.g., chorionic

23

villus cells or confluent cultured cells, the amount of
the above mentioned buffer with proteinase K may vary
according to the size of the tissue sample. The extract
is incubated for 4-10 hrs at 50°-60° C and then at 95° C
5   for 10 minutes to inactivate the proteinase. During
longer incubations, fresh proteinase K should be added
after about 4 hr at the original concentration.

When the sample contains a small number of cells,
extraction may be accomplished by methods as described
10   in Higuchi, "Simple and Rapid Preparation of Samples for
PCR", in PCR Technology, Ehrlich, H. A. (ed.), Stockton
Press, New York, which is incorporated herein by
reference. PCR can be employed to amplify target regions
from chromosome 7 in very small numbers of cells (1000-
15   5000) derived from individual colonies from bone marrow
and peripheral blood cultures. The cells in the sample
are suspended in 20 µl of PCR lysis buffer (10 mM Tris-
HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.1 mg/ml gelatin,
0.45% NP40, 0.45% Tween 20) and frozen until use. When
20   PCR is to be performed, 0.6 µl of proteinase K (2 mg/ml)
is added to the cells in the PCR lysis buffer. The
sample is then heated to about 60° C and incubated for 1
hr. Digestion is stopped through inactivation of the
proteinase K by heating the samples to 95° C for 10 min
25   and then cooling on ice.

A relatively easy procedure for extracting DNA for
PCR is a salting out procedure adapted from the method
described by Miller et al., Nucleic Acids Res. 16:1215
(1988), which is incorporated herein by reference.
30   Mononuclear cells are separated on a Ficoll-Hypaque
gradient. The cells are resuspended in 3 ml of lysis
buffer (10 mM Tris-HCl, 400 mM NaCl, 2 mM Na_2 EDTA, pH
8.2). Fifty µl of a 20 mg/ml solution of proteinase K
and 150 µl of a 20% SDS solution are added to the cells
35   and then incubated at 37° C overnight. Rocking the tubes

24

during incubation will improve the digestion of the
sample. If the proteinase K digestion is incomplete
after overnight incubation (fragments are still
visible), an additional 50 µl of the 20 mg/ml proteinase
5   K solution is mixed in the solution and incubated for
another night at 37° C on a gently rocking or rotating
platform. Following adequate digestion, one ml of a 6M
NaCl solution is added to the sample and vigorously
mixed. The resulting solution is centrifuged for 15
10   minutes at 3000 rpm. The pellet contains the
precipitated cellular proteins, while the supernatant
contains the DNA. The supernatant is removed to a 15 ml
tube that contains 4 ml of isopropanol. The contents of
the tube are mixed gently until the water and the
15   alcohol phases have mixed and a white DNA precipitate
has formed. The DNA precipitate is removed and dipped in
a solution of 70% ethanol and gently mixed. The DNA
precipitate is removed from the ethanol and air-dried.
The precipitate is placed in distilled water and
20   dissolved.

Kits for the extraction of high-molecular weight
DNA for PCR include a Genomic Isolation Kit A.S.A.P.
(Boehringer Mannheim, Indianapolis, Ind.), Genomic DNA
Isolation System (GIBCO BRL, Gaithersburg, Md.), Elu-
25   Quik DNA Purification Kit (Schleicher & Schuell, Keene,
N.H.), DNA Extraction Kit (Stratagene, La Jolla,
Calif.), TurboGen Isolation Kit (Invitrogen, San Diego,
Calif.), and the like. Use of these kits according to
the manufacturer's instructions is generally acceptable
30   for purification of DNA prior to practicing the methods
of the present invention.

The concentration and purity of the extracted DNA
can be determined by spectrophotometric analysis of the
absorbance of a diluted aliquot at 260 nm and 280 nm.
35   After extraction of the DNA, PCR amplification may

25

proceed. The first step of each cycle of the PCR
involves the separation of the nucleic acid duplex
formed by the primer extension. Once the strands are
separated, the next step in PCR involves hybridizing the

5   separated strands with primers that flank the target
sequence. The primers are then extended to form
complementary copies of the target strands. For
successful PCR amplification, the primers are designed
so that the position at which each primer hybridizes

10   along a duplex sequence is such that an extension
product synthesized from one primer, when separated from
the template (complement), serves as a template for the
extension of the other primer. The cycle of
denaturation, hybridization, and extension is repeated

15   as many times as necessary to obtain the desired amount
of amplified nucleic acid.

In a particularly useful embodiment of PCR
amplification, strand separation is achieved by heating
the reaction to a sufficiently high temperature for an

20   sufficient time to cause the denaturation of the duplex
but not to cause an irreversible denaturation of the
polymerase (see U.S. Pat. No. 4,965,188, incorporated
herein by reference). Typical heat denaturation involves
temperatures ranging from about 80° C to 105° C for

25   times ranging from seconds to minutes. Strand
separation, however, can be accomplished by any suitable
denaturing method including physical, chemical, or
enzymatic means. Strand separation may be induced by a
helicase, for example, or an enzyme capable of

30   exhibiting helicase activity. For example, the enzyme
RecA has helicase activity in the presence of ATP. The
reaction conditions suitable for strand separation by
helicases are known in the art (see Kuhn Hoffman-
Berling, 1978, CSH-Quantitative Biology, 43:63-67; and

35   Radding, 1982, Ann. Rev. Genetics 16:405-436, each of

26

which is incorporated herein by reference).

Template-dependent extension of primers in PCR is catalyzed by a polymerizing agent in the presence of adequate amounts of four deoxyribonucleotide
5 triphosphates (typically dATP, dGTP, dCTP, and dTTP) in a reaction medium comprised of the appropriate salts, metal cations, and pH buffering systems. Suitable polymerizing agents are enzymes known to catalyze template-dependent DNA synthesis.   In some cases,
10 the target regions may encode at least a portion of a protein expressed by the cell. In this instance, mRNA may be used for amplification of the target region. Alternatively, PCR can be used to generate a cDNA library from RNA for further amplification, the initial
15 template for primer extension is RNA. Polymerizing agents suitable for synthesizing a complementary, copy-DNA (cDNA) sequence from the RNA template are reverse transcriptase (RT), such as avian myeloblastosis virus RT, Moloney murine leukemia virus RT, or Thermus
20 thermophilus (Tth) DNA polymerase, a thermostable DNA polymerase with reverse transcriptase activity marketed by Perkin Elmer Cetus, Inc. Typically, the genomic RNA template is heat degraded during the first denaturation step after the initial reverse transcription step
25 leaving only DNA template. Suitable polymerases for use with a DNA template include, for example, E. coli DNA polymerase I or its Klenow fragment, T4 DNA polymerase, Tth polymerase, and Taq polymerase, a heat-stable DNA polymerase isolated from Thermus aquaticus and
30 commercially available from Perkin Elmer Cetus, Inc. The latter enzyme is widely used in the amplification and sequencing of nucleic acids. The reaction conditions for using Taq polymerase are known in the art and are described in Gelfand, 1989, PCR Technology, supra.
35

27

4. Allele Specific PCR

Allele-specific PCR differentiates between chromosome 7 target regions differing in the presence or absence of a variation or polymorphism. PCR amplification primers are chosen which bind only to certain alleles of the target sequence. Thus, for example, amplification products are generated from those chromosome 7 sets which contain the primer binding sequence, and no amplification products are generated in chromosome 7 sets without the primer binding sequence. This method is described by Gibbs, Nucleic Acid Res. 17:12427-2448 (1989).

5. Allele Specific Oligonucleotide Screening Methods

Further diagnostic screening methods employ the allele-specific oligonucleotide (ASO) screening methods, as described by Saiki et al., Nature 324:163-166 (1986). Oligonucleotides with one or more base pair mismatches are generated for any particular allele. ASO screening methods detect mismatches between variant target genomic or PCR amplified DNA and non-mutant oligonucleotides, showing decreased binding of the oligonucleotide relative to a mutant oligonucleotide. Oligonucleotide probes can be designed that under low stringency will bind to both polymorphic forms of the allele, but which at higher stringency, bind to the allele to which they correspond. Alternatively, stringency conditions can be devised in which an essentially binary response is obtained, i.e., an ASO corresponding to a variant form of the CYP11a1 gene will hybridize to that allele, and not to the wildtype allele.

28

6. Ligase Mediated Allele Detection Method

     Target regions of a test subject's DNA can be
compared with target regions in unaffected and affected
family members by ligase-mediated allele detection. See
Landegren et al., Science 241:1077-1080 (1988). Ligase
may also be used to detect point mutations in the
ligation amplification reaction described in Wu et al.,
Genomics 4:560-569 (1989). The ligation amplification
reaction (LAR) utilizes amplification of specific DNA
sequence using sequential rounds of template dependent
ligation as described in Wu, supra, and Barany, Proc.
Nat. Acad. Sci. 88:189-193 (1990).

7. Denaturing Gradient Gel Electrophoresis

     Amplification products generated using the
polymerase chain reaction can be analyzed by the use of
denaturing gradient gel electrophoresis. Different
alleles can be identified based on the different
sequence-dependent melting properties and
electrophoretic migration of DNA in solution. DNA
molecules melt in segments, termed melting domains,
under conditions of increased temperature or
denaturation. Each melting domain melts cooperatively at
a distinct, base-specific melting temperature (Tm).
Melting domains are at least 20 base pairs in length,
and may be up to several hundred base pairs in length.
     Differentiation between alleles based on sequence
specific melting domain differences can be assessed
using polyacrylamide gel electrophoresis, as described
in Chapter 7 of Erlich, ed., PCR Technology, Principles
and Applications for DNA Amplification, W.H. Freeman and
Co, New York (1992), the contents of which are hereby
incorporated by reference.

29

Generally, a target region to be analyzed by
denaturing gradient gel electrophoresis is amplified
using PCR primers flanking the target region. The
amplified PCR product is applied to a polyacrylamide gel
5   with a linear denaturing gradient as described in Myers
et al., Meth. Enzymol. 155:501-527 (1986), and Myers et
al., in Genomic Analysis, A Practical Approach, K.
Davies Ed. IRL Press Limited, Oxford, pp. 95-139 (1988),
the contents of which are hereby incorporated by
10   reference. The electrophoresis system is maintained at a
temperature slightly below the Tm of the melting domains
of the target sequences.

In an alternative method of denaturing gradient gel
electrophoresis, the target sequences may be initially
15   attached to a stretch of GC nucleotides, termed a GC
clamp, as described in Chapter 7 of Erlich, supra.
Preferably, at least 80% of the nucleotides in the GC
clamp are either guanine or cytosine. Preferably, the GC
clamp is at least 30 bases long. This method is
20   particularly suited to target sequences with high Tm's.

Generally, the target region is amplified by the
polymerase chain reaction as described above. One of the
oligonucleotide PCR primers carries at its 5' end, the
GC clamp region, at least 30 bases of the GC rich
25   sequence, which is incorporated into the 5' end of the
target region during amplification. The resulting
amplified target region is run on an electrophoresis gel
under denaturing gradient conditions as described above.
DNA fragments differing by a single base change will
30   migrate through the gel to different positions, which
may be visualized by ethidium bromide staining.

8. Temperature Gradient Gel Electrophoresis

35   Temperature gradient gel electrophoresis (TGGE) is

30

based on the same underlying principles as denaturing
gradient gel electrophoresis, except the denaturing
gradient is produced by differences in temperature
instead of differences in the concentration of a
5.   chemical denaturant. Standard TGGE utilizes an
electrophoresis apparatus with a temperature gradient
running along the electrophoresis path. As samples
migrate through a gel with a uniform concentration of a
chemical denaturant, they encounter increasing
10   temperatures. An alternative method of TGGE, temporal
temperature gradient gel electrophoresis (TTGE or tTGGE)
uses a steadily increasing temperature of the entire
electrophoresis gel to achieve the same result. As the
samples migrate through the gel the temperature of the
15   entire gel increases, leading the samples to encounter
increasing temperature as they migrate through the gel.
Preparation of samples, including PCR amplification with
incorporation of a GC clamp, and visualization of
products are the same as for denaturing gradient gel
20   electrophoresis.

9. Single-Strand Conformation Polymorphism Analysis

Target sequences or alleles at the CYP11a1 locus
25   can be differentiated using single-strand conformation
polymorphism analysis, which identifies base differences
by alteration in electrophoretic migration of single
stranded PCR products, as described in Orita et al.,
Proc. Nat. Acad. Sci. 86:2766-2770 (1989). Amplified PCR
30   products can be generated as described above, and heated
or otherwise denatured, to form single stranded
amplification products. Single-stranded nucleic acids
may refold or form secondary structures which are
partially dependent on the base sequence. Thus,
35   electrophoretic mobility of single-stranded

31

amplification products can detect base-sequence
difference between alleles or target sequences.

10. Chemical or Enzymatic Cleavage of Mismatches
        Differences between target sequences can also be
detected by differential chemical cleavage of mismatched
base pairs, as described in Grompe et al., Am. J. Hum.
Genet. 48:212-222 (1991). In another method, differences
between target sequences can be detected by enzymatic
cleavage of mismatched base pairs, as described in
Nelson et al., Nature Genetics 4:11-18 (1993). Briefly,
genetic material from a patient and an affected family
member may be used to generate mismatch free
heterohybrid DNA duplexes. As used herein,
"heterohybrid" means a DNA duplex strand comprising one
strand of DNA from one person, usually the patient, and
a second DNA strand from another person, usually an
affected or unaffected family member. Positive selection
for heterohybrids free of mismatches allows
determination of small insertions, deletions or other
polymorphisms that may be associated with alterations in
androgen metabolism.

11. Non-PCR Based DNA Diagnostics
        The identification of a DNA sequence linked to
CYP11a1 can be made without an amplification step, based
on polymorphisms including restriction fragment length
polymorphisms in a patient and a family member.
Hybridization probes are generally oligonucleotides
which bind through complementary base pairing to all or
part of a target nucleic acid. Probes typically bind
target sequences lacking complete complementarity with
the probe sequence depending on the stringency of the
hybridization conditions. The probes are preferably
labeled directly or indirectly, such that by assaying

32

for the presence or absence of the probe, one can detect
the presence or absence of the target sequence. Direct
labeling methods include radioisotope labeling, such as
with ^{32}P or ^{35}S. Indirect labeling methods include
fluorescent tags, biotin complexes which may be bound to
avidin or streptavidin, or peptide or protein tags.
Visual detection methods include photoluminescents,
Texas red, rhodamine and its derivatives, red leuco dye
and 3, 3', 5, 5'-tetramethylbenzidine (TMB),
fluorescein, and its derivatives, dansyl, umbelliferone
and the like or with horse radish peroxidase, alkaline
phosphatase and the like.

Hybridization probes include any nucleotide
sequence capable of hybridizing to the porcine
chromosome where CYP11a1 resides, and thus defining a
genetic marker linked to CYP11a1, including a
restriction fragment length polymorphism, a
hypervariable region, repetitive element, or a variable
number tandem repeat. Hybridization probes can be any
gene or a suitable analog. Further suitable
hybridization probes include exon fragments or portions
of cDNAs or genes known to map to the relevant region of
the chromosome.

Preferred tandem repeat hybridization probes for
use according to the present invention are those that
recognize a small number of fragments at a specific
locus at high stringency hybridization conditions, or
that recognize a larger number of fragments at that
locus when the stringency conditions are lowered.

The following examples are provided to illustrate
embodiments of the present invention. They are not
intended to limit the invention in any way.

33

EXAMPLE I

A Genetic Marker for Meat Quality, Growth, Carcass and
Reproductive Traits in Pigs

5   In accordance with the present invention, a genetic
marker has been identified and characterized which is
associated with improved meat quality and improved
growth and carcass traits in pigs. The following
materials and methods were utilized in the practice of
10   Example I.

Testis tissue samples were obtained from fifty
Yorkshire boars for which growth, carcass, and boar
taint data had previously been collected. Boars were
weaned at approximately 10 weeks of age, assigned to
15   pens, and fed a standard grower-finisher diet to a final
weight of approximately 120 kg. Boars were killed by
electrical stunning and exsanguination at the Penn State
University meats Laboratory. Testes, bulbourethral
glands and submaxillary salivary glands were collected,
20   trimmed, and weighed. Carcasses were weighed and then
chilled overnight. The following day data were collected
for standard carcass measurements such as carcass
length, loin eye area, fat depth and marbling.

The assay for submaxillary salivary gland delta-16-
25   androstenes was adapted from a procedure developed by
Squires (J. Animal Sci. 69: 1092-1100, 1991). Briefly,
submaxillary salivary glands were trimmed and minced in
a food processor (Cusinart) and one gram of minced
tissue was placed in a 50 ml test tube. Methanol (5 ml)
30   was added and the mixture was homogenized for 30 sec by
Polytron. Samples were placed in a centrifuge for 5 min
@ 2800 rpm. Three ml of distilled water were added to 3
ml of the supernatant and mixed by vortexing. Six ml of
hexane were added to extract the delta-16-androstenes.
35   The mixture was vortexed and allowed to stand for 5 min

34

for the phases to separate. Three milliliters of the
organic phase were transferred to a glass culture tube
and the extract was dried under nitrogen while in a
water bath (30°C). Color reagents were added (.5 ml of
.5% resorcylaldehyde in glacial acetic acid plus .5 ml
of 5% sulfuric acid in glacial acetic acid). The tubes
were placed in a heat block for 10 min at 95 C.
Development of a violet color, an index of the presence
of delta-16-androstenes, was measured by pipetting 100
μl of the test solution into a well in a 96-well
microplate. Absorbance was measured at several
wavelengths near the known absorbance maximum for Δ16-
androstenes (593 nm) and compared against standard test
solutions containing 5α-androst-16-ene-3β-ol (the
predominant Δ16-androstene in the submaxillary salivary
gland). Concentration of Δ16-androstenes was established
by generation of a standard curve with the standard test
solutions.

Data were analyzed by ANOVA using the GLM
procedures of SAS (1990).

Testis tissue samples were obtained from storage
(-20°C) for ten boars: five that had the highest
concentrations of Δ16-androstenes (high boar taint) and
five that had the lowest concentrations of Δ16-
androstenes (low boar taint). DNA was extracted by
Proteinase K digestion. Approximately 50 mg of testis
tissue was wrapped in aluminum foil and frozen in liquid
nitrogen. The sample was then pulverized and
approximately 20 mg was placed in a microfuge tube with
.5 ml digestion buffer (50 mM Tris, pH 8.5; 1mM EDTA;
0.5% Tween 20; 200 μg/ml proteinase K (Gibco Life
Technologlies, Grand Island, NY). Proteinase K was
stored at -20°C in stock solution (20 mg/ml proteinase
K; 1-mM Tris-HCl, pH 7.5; 20 mM calcium chloride, and 5%
glycerol). The samples were suspended in digestion

35

buffer and placed in a water bath @ 55°C for 3 hours.
Samples were centrifuged for 1 min @13,000 g and placed
in a heat block for 10 min @ 95°C. Samples were removed
and stored at -20°C until analyzed.

5   Four sets of primers were obtained which
corresponded to approximately 600 bp each for a total of
approximately 2.4kb of the 5'UTR of the porcine CYP11a1
gene (sequence obtained from Urban, et al., J. Biol.
Chem. 269:25761-25769, 1994). See Figure 1. Polymerase

10   Chain Reactions were initiated for each primer set for
each of the ten DNA templates. PCR was performed as
follows.

        1. 2 min @ 94 C.
        2. 1 min @ 94 C

15   3. 1 min @ 55 C
        4. 1 min @ 72 C
        5. 35 cycles to (2.)
        6. 2 min @ 72 C
        7. hold at 5 C

20

        Reactions were performed using 10x buffer (w/MgCl$_2$);
dNTP's (10 nmol); primer CYPscc For1 (20 pmol);primer
CYPscc Rev1 (20 pmol); Taq polymerase ;ddH$_2$O and DNA
template (1:10 dilution of Proteinase K digested sample,

25   approximately 100 ng).
        PCR products were analyzed by agarose gel
electrophoresis, and the ~600 bp bands cut out of the
agarose gel and purified using the QIAquick gel
extraction kits (QIAGEN Inc., Valencia CA). The

30   nucleotide sequences of each of the forty PCR products
was determined in both forward and reverse directions
using an ABI Prism Model 377 Sequencer (Perkin Elmer,
CA) at the Penn State Nucleic Acid Facility, PSU
Biotechnology Institute.

36

The sequences of the PCR products were aligned
manually and examined for differences between the ten
animals. While there were 37 differences in the samples
when compared with the published sequence (Urban et al.,
5   1994, *supra*), there was only one base pair that varied
among this group of animals. At position -155 (155 bases
before the start site ATG codon), six of the samples had
the cytosine (CC), and four were polymorphic; that is
they had both the cytosine and thymidine (CT),
10   indicating heterozygosity at that base pair. Of
significant interest was that all five of the high taint
boar samples were the CC genotype, whereas four out of
five of the low taint boar samples had the CT genotype.
This polymorphism was located in a restriction
15   enzyme recognition site such that the presence of the T
allele would change the restriction fragment length
pattern observed after digestion with specific
restriction enzymes. In this particular case, the
restriction enzyme used was SphI (New England Biolabs).
20   Presence or absence of the T allele in the DNA samples
from the full group of fifty boars was determined by
Restriction Fragment Length Polymorphism analysis
involving examination of restriction digests of CYP11a1
5'UTR using SphI. For exemplary gel, see Figure 2.
25   Presence of the T allele, either homozygous (TT) or
heterozygous (CT) was associated with low boar taint.
Presence of the CC allele was associated with high boar
taint, as well as with increased testis weight,
increased bulbolurethral gland length and weight, and
30   increased submaxillary salivary gland weight. See
Figure 3 and Table 4. In addition, boars that possessed
the CC allele exhibited a 5.9% improvement in rate of
gain, and had greater amounts of lean muscle as
evidenced by longer carcasses, and tended to have less
35   fat as determined by backfat depth measurements. Boars

37

with the CC allele also tended to have higher
concentrations of serum testosterone in blood samples
taken at slaughter.

     A retrospective analysis of production records of
direct female relatives (dams and siblings) of these
boars revealed that those females related to boars
possessing the T allele tended to have slightly larger
litter sizes (+.31 pigs/litter) and weaned heavier
litters (+4.27 kg). Thus this polymorphism appears to
confer·beneficial fertility and productivity traits to
female pigs.

EXAMPLE II
A Genetic Marker for Meat Quality, Growth, Carcass and
Reproductive Traits in Cows and Chickens

     The identification and characterization of the
CYP11a1 polymorphism in pigs facilitates the
characterization of the corresponding polymorphism in
bovines which are associated with improved reproductive
and carcass traits. The bovine CYP11a1 sequence is
provided in Figure 5. A suitable primer set for
amplifying the bovine homologue of the 5' UTR for the
CYP11a1 gene has the following sequences: Sense:
5'-GCAGATGTCCCTGGTGATTC-3'; and Antisense:
5'-TGAACGGAGGGGAAGCC-3'.

     Amplified bovine CYP11a1 sequences and
corresponding genetic traits are then characterized as
set forth herein for the porcine CYP11a1 gene.

     Figure 6 depicts the CYP11a1 gene from chicken. In
order to assess genetic changes in a more lengthy 5'UTR
sequence from the chicken CYP11a1 sequence provided in
Genbank, the cDNA sequence provided in Figure 6 is
utilized as the basis for 5' rapid amplification of cDNA
ends (RACE) using a kit from Clontech containing RACE-

38

ready cDNA prepared from chicken. Clones obtained from
this RACE approach yield 5' end points of the chicken
CYP11a1 sequence for further analysis of genetic changes
in the 5'UTR associated with improved reproductive and
5   carcass traits. Genetic polymorphisms and alterations
so identified are within the scope of the present
invention. Suitable protocols for practicing RACE are
provided in Current Protocols of Molecular Biology, J.
Wiley & Sons, Inc. 1998, Chapter 15.6.9, the entire
10   disclosure of which is incorporated by reference herein.

The present invention is not limited to the
embodiments specifically described above, but is capable
of variation and modification without departure from the
15   scope of the appended claims.

39

Claims

5

10

15

20

25

30

35

40

45

50

55

What is claimed is:

1. A method of screening test subjects to identify those more likely to have better growth, development, reproduction and carcass traits such as rates of gain, carcass length, or litter size, comprising: obtaining a sample of genetic material from a test subject and assaying for the presence of a polymorphism in the CYP11a1 gene which is associated with rate of gain, carcass length, and litter size.

2. The method of claim 1 wherein said step of assaying is selected from the group consisting of restriction fragment length polymorphism (RFLP) analysis, heteroduplex analysis, single strand conformational polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE).

3. The method of claim 1, wherein said step of assaying for the presence of said polymorphism comprises the steps of digesting said genetic material with a restriction enzyme that cleaves the CYP11a1 gene in at least one place; separating the fragments obtained from the said digestion; detecting a restriction pattern generated by said fragments; and comparing said pattern with a second restriction pattern for the CYP11a1 gene obtained by using said restriction enzyme, wherein said second restriction pattern is associated with increased rates of gain, increased carcass length, and increased litter size.

4. A method as claimed in claim 1, wherein said test subject is selected from the group consisting of pigs, cows and chickens.

40

5. The method of claim 3 wherein said restriction
enzyme is SphI and said test subject is a pig.

6. The method of claim 3 wherein said separation is
by gel electrophoresis.

7. The method of claim 3 wherein said step of
comparing said restriction patterns comprises
identifying specific fragments by size and comparing the
sizes of said fragments.

8. The method of claim 5 further comprising the
step of amplifying the amount of porcine CYP11a1 gene or
a portion thereof which contains said polymorphism,
prior to said digestion step.

9. The method of claim 3 wherein said restriction
site is located in the untranslated region of the
CYP11a1 gene.

10. The method of claim 7 wherein said
amplification includes the steps of selecting a forward
and a reverse sequence primer capable of amplifying a
region of the porcine CYP11a1 gene which contains a
polymorphic restriction site.

11. The method of claim 10 wherein said forward and
reverse primers are between 10 and 50 nucleotides in
length and selected from SEQ ID NO: 1.

12. The method of claim 10 wherein said forward
primer is SEQ ID NO:2 and said reverse primer is SEQ ID
NO:3.

13. The method of claim 6 wherein said step of

41

detecting sizes of said fragments comprises the steps of
separating said fragments by size using gel
electrophoresis in the presence of a control DNA
fragment of known size; contacting said separated
5   fragments with a probe that hybridizes with said
fragments to form probe-fragment complexes; and
determining the size of separated fragments by detecting
the presence of the probe fragment.

10   14. A method for identifying a genetic marker for
pig growth rate, carcass length, litter size, or boar
taint comprising the steps of breeding male and female
pigs of the same breed or breed cross or derived from
similar genetic lineages; determining the growth rates,
15   carcass lengths, number of offspring, or presence of
boar taint; determining the presence of a polymorphism
in the CYP11a1 gene of each pig; and associating the
growth rate, carcass length, number of offspring, or
presence of boar taint of each pig with said
20   polymorphism thereby identifying a polymorphism for
these traits.

         15. The method of claim 14 further comprising the
step of selecting pigs for breeding which are predicted
25   to have better growth rates, longer carcasses, increased
litter size, or decreased boar taint by said marker.

         16. The method of claim 14 wherein said analysis
comprises digestion of PCR amplified DNA with the
30   restriction enzyme Sphl.

         17. The method of claim 12 wherein said
polymorphism associated with growth rate, carcass
length, litter size, or boar taint is detected by use of
35   forward and reverse primers comprising at least 4

                                42

consecutive bases in SEQ NOS: 2 and 3.

18. A kit for evaluating a sample of porcine DNA
comprising, in a container, a reagent that identifies a
5   polymorphism in the porcine CYP11a1 gene.

19. The kit of claim 18 wherein said reagent is a
primer that amplifies the porcine CYP11a1 gene or a
fragment thereof.
10

20. The kit of claim 18 further comprising a DNA
polymerase, a restriction enzyme which cleaves the
porcine CYP11a1 gene in a least one place; and forward
and reverse primers capable of amplifying a region of
15   the porcine CYP11a1 gene which contains a polymorphic
site.

21. A primer for assaying for the presence of a
polymorphic Sph1 site in the porcine CYP11a1 gene
20   wherein said primer comprises a sequence from the group
of SEQ ID NO:2 and SEQ ID NO:3.

22. A genetic marker associated with growth rate,
carcass length, litter size, and boar taint in pigs,
25   said marker comprising a polymorphism in the porcine
CYP11a1 gene.

23. The genetic marker of claim 22 wherein said
polymorphism is a Sph1 restriction site.
30

24. The marker of claim 22 wherein said
polymorphism is located in the 5' untranslated region of
the porcine CYP11a1 gene.

35   25. A DNA sequence from the porcine CYP11a1 gene 5'

43

untranslated region, said sequence consisting of SEQ ID
NO: 1.

26. A primer designed to amplify a polymorphic Sph1
restriction site in the porcine CYP11a1 gene wherein
said primer is 4 or more continuous bases from SEQ ID
NO: 1.

27. A primer designed to amplify a polymorphic Sph1
restriction site in the porcine CYP11a1 gene wherein
said primer is a reverse primer generated from the SEQ
ID NO: 1.

28. A method for screening pigs to determine those
more likely have increased growth rates, longer
carcasses, larger litters, higher boar taint, and/or
those less likely to exhibit increased growth rates,
longer carcasses, larger litters, or higher boar taint,
which method comprises of the steps: determining the
alleles of the CYP11a1 gene present in a pig;
determining the alleles of other markers for genes know
to affect growth rate, carcass length, litter size, or
boar taint; and selecting for animals with favorable
combinations of alleles and against those carrying
unfavorable combinations.

29. The method of claim 28 wherein the
determination of CYP11a1 alleles comprises determining
the presence of at least one allele associated with at
least one DNA marker linked either directly or
indirectly to CYP11a1.

30. The method of claim 28 wherein the DNA marker
is a microsatellite.

44

31. The method of claim 28 wherein the DNA marker is SO064, SO102, S0078, SO158, S0066, SW304, SW1083, SO101, or S0212.

32. The method of claim 28 wherein the marker is selected from the group of tumor necrosis factor alpha (TNFα), CYP11a1, prolactin (PRL), estrogen receptor (ER) and prolactin receptor (PRLR).

45

Figure 1.

SEQ ID NO:1

GCTCCAAAGAGACATTTTGGGGTGGCAAAATAGTCTACAGGATTCTATGGCATA
GGAGACAACTCTCAGATAGCTCTGCAGACCTGCTCCAAAGAAGTATAGGAGAAG
CCAGGATTTATAAGAACTTTTTTGTTGGGAAAATAAATGTAGTCAAACATAAAAAG
ACAACTGCTAATAACAAACAATAGACATGTCAAGATAATGACCTTAGTGCCTTTCT
ATGTGTGGAAAGACTCAAGAATCTGGGGTCATTGAACTTTTTCCTTAGATATGCA
TCTTAATATCCTGGGGTCAGTATAATCCAAATGCTTCCTGTTTTTCTCCATCCTAA
AGTCCCCTCCGGGTGCACTGATGGGTTCCCCTCCAGTGGGCAACTGCAGTGGC
AATTGGCTTGATCTCTGTAGAACTGGAATGGTGGGCAACATTCTTTTCTTTACAG
TATCCTGAGTCTGGGAGGGGCTGTGTGGGCCAGAGCCTGNATGCAGGAGGAG
GAGGGAGTCTGATCGCTTAGTCAGCTTCTCGCTTAACCTTGAGCTGGTGGTTAT
AAGCTGGGCCCCAGGCGCCCGAGGCCAGACTCACCTCATCAGGCCCTGCTGCA
GTGGGAGCAGGGAGAGTAGCAGTGGTAGGGGCAGCATG

N = C or T at polymorphic site

SEQ ID NO:2

Forward primer:

GCTCCAAAGAGACATTTTGGGGTGGC

SEQ ID NO:3
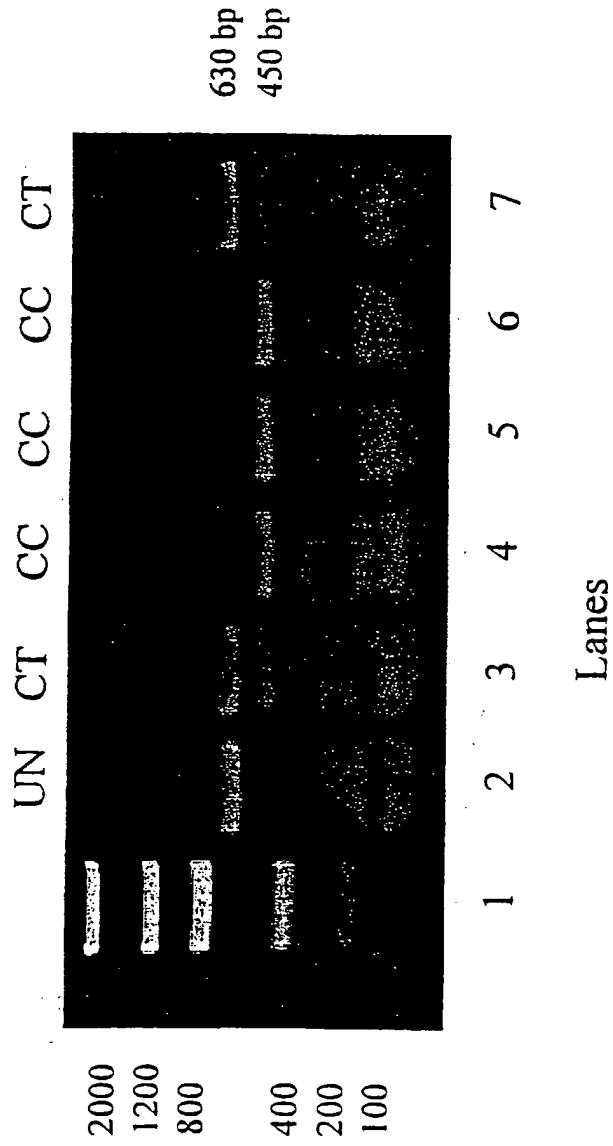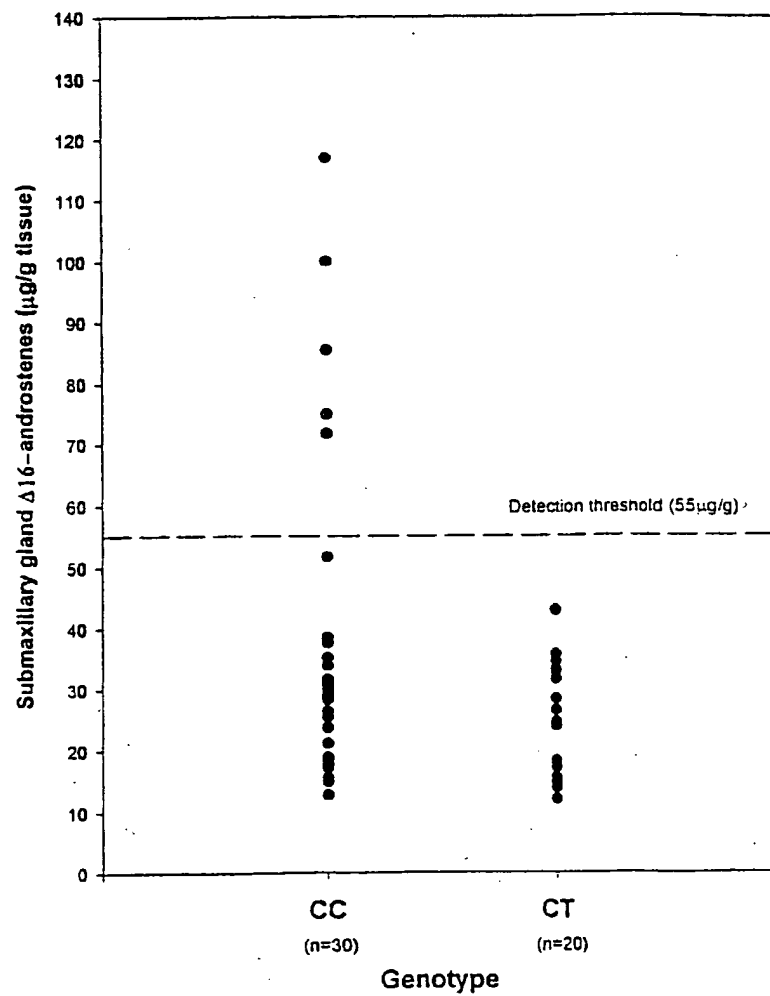
Reverse primer:

CATGCTGCCCCTACCACTGCTACTCT

Figure 2. SphI restriction digest of porcine CYP11a1 PCR fragment

Figure 3. Comparison of submaxillary salivary gland Δ16-androstenes
in boars possessing a CYP11a1 single nucleotide polymorphism.

Genotype

	CC	CT	P value
Rate of gain (kg BW/d)	0.76 ± .01	0.72 ± .01	.05
Carcass length (cm)	85.17 ± .38	82.96 ± .47	.001
Submaxillary salivary gland (SMG) wt (g)	92.1 ± 3.1	71.5 ± 4.9	.0001
Δ16-androstenes in SMG (μg/g)	38.7 ± 4.1	23.9 ± 5.0	.05
Relative SMG wt (g/kg BW)	0.72 ± .023	0.58 ± .027	.001
Bulbourethral gland length (mm)	128.8 ± 2.4	117.7 ± 2.9	.01
Relative bulbourethral gland wt (g/kg BW)	93.8 ± 4.0	73.5 ± 4.9	.01
Testis wt (g)	628.6 ± 27.1	530.2 ± 25.4	.05
Relative testis wt (g/kg BW)	4.92 ± .20	4.33 ± .24	.10
Serum testosterone at slaughter (ng/ml)	2.04 ± .28	1.59 ± .35	.32

Figure 4. Growth, carcass, and reproductive traits of pigs with CC or CT
CYP11a1 polymorphism.

```
   1 gcagatgtcc ctggtgattc ctgaaacagg ccctctgttt aaattcttca gcagttagag
  61 ggaaggtcaa tttttcccaa ggcttttggg ctttgattgt tttcattttt aaattatctg
 121 cattctaaag agatattttg ggtggcagat tttgctctcc tacaggactt tgtctaggag
 181 acggctctca ggccagctcc gacgactgtt ccaaagaagt aagggaaagc tagggtttat
 241 atcaatcttt tttttttgctg ggagaagggg gatgaacatg tagtcaaaca taaaaagatc
 301 actgctaatc ccaaacaaca gacacctcaa gtgaatggtt ttagtgtttt tctatatatg
 361 ttgtttagtc actaagtcct gtccgactct tttgcgactc catagactgt agcccaccaa
 421 gctcctctgt ccatgggatt tttctaggca agaatactgg agtgggttgc catttccttc
 481 tccctgggat cttcctaacc caaggactga acccttgtct cctgcattgc aggtggattt
 541 tttaccgact gagccaccag ggaagttatg tgtgcaagaa tccggggtca tcgaaatttt
 601 cccttagata tacatcgtat ctagggacca gtacaatgca aatgcttcct gttttctttc
 661 atcctgaagt ctcctcaggg tgcattgagg gagggagtcc cctcaggtgg gtgaccacag
 721 tggctgacgc ttgatgttgt agaactggaa tgatgggtta cattctttcg tttacagtac
 781 tgagtctggg aggagctgtg tgggctggag tcagccggag gaggctgacc gccctgtcag
 841 cttctcactt agccttgagc tggtgattat aagctgggtc ccagggtccc agggccagag
 901 tcacctgctg cagtacgagc agagacagca gcagctgtgg gggcagcatg ctagcaaggg
 961 ggcttcccct ccgttcagcc ctggtcaaag cctgcccacc catcctgagc tcagtggggg
1021 agggctgggg ccaccacagg gtgggcactg gagagggagc tggcatctcc acaaagaccc
1081 ctcgcccta cagtgagatc ccctcccctg gtgacaatgg ctggcttaac ctctaccatt
1141 tctggaggga gaagggctca cagagaatcc actttcgcca catcgagaac ttccagaagt
1201 atggcccccat ttacaggtaa gcctggcagg aggattgggg ctggcgggat agggaagcct
1261 gtggtggccc cctcctgaa aggtctgccc tccccttcca ggctctggtt cacctctgac
1321 tttatttctt cctgcctggc ggtggcagga gtagagttaa tgcttccag acagtgggtt
1381 cacttcccag ccctgaggcc tcaacagtcc ccgggctcta caccctttaga aactttgggg
1441 aggtgggggag gcccaagaaa ataagccccg g
```

FIGURE 5

```
   1 cttttttcgg ttgtaccttt gtctctgtac agatattttg taatatatta aaaacaaaac
  61 ctactgagct cctcgccttg agcccaggat tcagggataa gagcgaggtc gccccggccg
 121 tgcgccgccc tgctcccatg ctctccaggg ctgcacccat agcgggcagc tttcaggcat
 181 gccgctgtgc cggagggatc ccagccctcg cggggggtcca ctacccattg cccagctcct
 241 cgggagctcg gcctttcgac caggtgccgg gtgaatggag agcgggttgg ctcaacctgt
 301 accacttctg gaaggaggga ggcttccaca acgtgcacaa catcatggcc agcaagttcc
 361 agcgctttgg gcccatctac agggagaagt tgggtgtcta cgagagcgtg aatatcatca
 421 gccccgcgca tgcggccacg ctcttcaagt cagaggggat gctgcccgag cgcttcagcg
 481 tgcccccatg ggtggcatac cgtgactacc gcaacaagcc ctacggcgtg ctcctcaaga
 541 caggggaggc ctggccgctcg gaccgcctga ccctgaacaa ggaggtgctg tcgccgcagg
 601 tggtggacag cttcgtgccc ttgctggacc aggtgagcca ggactttttg cggcgggcac
 661 gggcgcaggt ccagcagagc ggccgggagc gctggacggc cgacttcagc cacgagctct
 721 tccgctttgc cttggagtct gtgtgccacg tgctgtatgg gaacgcctg gggctgctgc
 781 aggactttgt ggacccagag gcacagcagt tcatcgacgc cgtcaccctc atgttccaca
 841 ccacctcccc catgctctac gtgccacccg ccctgctccg ccacctcaac accaagacat
 901 ggcgtgacca cgtgcatgct tgggatgcca tcttcacaca ggctgacaaa tgtatccaaa
 961 acgtttaccg ggacatccgg ctgcaacgca agagcaccga ggagcacacg ggcatcctct
1021 tcagcctcct tgtgcaggac aagctgcccc tggatgacat caaggccagc gtcaccgaga'
1081 tgatggcggg cggcgtggac acgacttcca tgactctgca atgggccatg ctggagctgg
1141 cacgatcccc gggcatccag gagcggctgc gggcagaggt gctggcagcc aagcaggagg
1201 cacaggggga cagggtgaag atgctgaaga gcatccgact gctcaaagcc gccatcaagg
1261 agactctcag gctgcacccg gtggcggtga cgctgcagag gtacaccaca caggaggtca
1321 tcctgcagga ctaccgcatc cccccaaga cgctggtgca ggttggtctc tacgccatgg
1381 gacgagaccc tgaggtcttc cccaagccgg agcagttcaa ccctgagcgc tggctggtga
1441 tgggctccaa gcacttcaag ggactgagct ttgggtttgg gccacggcag tgtctgggtc
1501 gtcgcatcgc cgagctggag atgcagctct tcctcatgca catcctggag aactttaaga
1561 tcgaaaccaa gcgggcggtg gaagttggga ccaagttcga cctcattctt gtccctgaaa
1621 aacccatcta cctgagactg cggcccctcc agcccagga gtgacatggg gtgtccccag
1681 ttggtcccag cttggggaca cctccatcag ctcagcgcat tcagccttgg ctccagccct
1741 tcttacgcca tgggggagat ggctgcccccc ttcccatttt cttcgcctct gatttgctct
1801 gtaatttctg caccaaaagc
```

FIGURE 6

INTERNATIONAL SEARCH REPORT

A.   CLASSIFICATION OF SUBJECT MATTER

IPC(7)   :C07H 21/04; C12Q 1/68
US CL   :435/6; 536/23.1, 24.33

According to International Patent Classification (IPC) or to both national classification and IPC

B.   FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :   435/6; 536/23.1, 24.33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C.   DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 99/18192 A1 (THE PENN STATE RESEARCH FOUNDATION) 15 April 1999, see entire document.	1-32
A	NOLAN et al. Genotype of the P450scc locus determines differences in the amount of P450scc protein and maximal testosterone production in mouse Leydig cells. Mol. Endocrinol. October 1990. Vol. 4, No. 10, pages 1459-1464, see entire document.	1-32
A	DUROCHER et al. Genetic linkage mapping of the CYP11a1 gene encoding the cholesterol side-chain cleavage P450scc close to the CYP1a1 gene and D15S204 in the chromosome 15q22.33-q23 region. Pharamacogenetics. February 1998, Vol. 8, No. 1, pages 49-53, see entire document.	1-32

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	
"E" earlier document published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 AUGUST 2000	05 SEP 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer ROBERT SCHWARTZMAN
Facsimile No. (703) 305-3230	Telephone No. (703) 305-0196

Form PCT/ISA/210 (second sheet) (July 1998) *

INTERNATIONAL SEARCH REPORT

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	DAVIS et al. Association of cytochrome b5 with 16-androstene steroid synthesis in the testis and accumulatioin in the fat of male pigs. J. Anim. Sci. May 1999, Vol. 77, No. 5, pages 1230-1235, see entire document.	1-32

INTERNATIONAL SEARCH REPORT

Internal application No.
PCT/US00/13168

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
   because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 11, 12, 17, 21, 25-27
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

   The claims are drawn to specific SEQ ID NOS but the claims could not be searched as the computer readable form of the Sequence Listing filed July 27, 2000 was blank (See attached CRF Problem Report).

3. ☐ Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1)) (July 1998) *

INTERNATIONAL SEARCH REPORT

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

STN: Medline Biosis Embase CAPlus
WEST
Search Terms: boar taint, boar odor, CYP11a1, cytochrome P450, polymorphism